Simulating episodic memory deficits in semantic dementia with the TraceLink model.

Meeter, M.; Murre, J.M.J.

published in
Memory
2004

DOI (link to publisher)
10.1080/09658210244000658

document version
Publisher's PDF, also known as Version of record

MEMORY, 2004, *12* (3), 272–287

Simulating episodic memory deficits in semantic dementia with the TraceLink model

Martijn Meeter and Jaap M. J. Murre

University of Amsterdam, The Netherlands

Although semantic dementia is primarily characterised by deficits in semantic memory, episodic memory is also impaired. Patients show poor recall of old autobiographical and semantic memories, with better retrieval of recent experiences; they can form new memories, and normal performance on pictorial recognition memory has been demonstrated. As these abnormalities in episodic memory are virtually a mirror image of those seen in the amnesic syndromes, semantic dementia poses a challenge to extant models of remote memory and amnesia. Here, we show that one such model, TraceLink, can reproduce some of the principal findings on episodic memory in semantic dementia. A loss of nodes and connections within the trace system, which can be identified with the temporal neocortical memory storage sites implicated in semantic dementia, simulates without further assumptions the findings reported above.

INTRODUCTION

In order to accommodate conflicting and complex findings from amnesia (Bauer, Tobias, & Valenstein, 1993; Kapur, 1999), a number of researchers have proposed that medial temporal lobe structures (i.e., the hippocampus, subiculum, parahippocampal gyrus and entorhinal cortex) play a temporary, time-limited role in the acquisition of human long-term memories (Marr, 1971; Milner, 1989; Squire & Alvarez, 1995; Squire, Cohen, & Nadel, 1984). In this line of thinking, the hippocampal complex is necessary for the retrieval of recently experienced events, but is not involved in the retrieval of older episodic and semantic memories. By contrast, regions of the neocortex are thought to be the more permanent repository of memory (Squire & Alvarez, 1995). This theory provides an explanation for why patients with damage to the hippocampal complex show a temporally graded loss of memory, with recent memories being affected more severely than older memories.

The verbal theory described above has inspired several neuroanatomically informed computational models of human memory. A number of researchers have shown that lesioning of the hippocampus in these models results in a temporal gradient in memory retrieval similar to that seen in amnesic patients with hippocampal damage (Alvarez & Squire, 1994; McClelland, McNaughton, & O'Reilly, 1995; Murre, 1996). Other experimental data related to retrograde amnesia, derived from studies with amnesic patients, have also been tested and modelled (Murre & Meeter, 2002). This approach to the study of human long-term memory allows neuropsychological theories to be tested in detail, and results in the generation of new hypotheses which can be investigated in brain-damaged subjects (for reviews of these computational models, see Murre, in press, and special issues of *Hippocampus*, 1996, issue 6, and *Memory*, 1997, issue Jan–March).

One condition that has not been investigated so far with these models, is semantic dementia, also called progressive fluent aphasia (Hodges,

Correspondence should be addressed to M. Meeter, Dept. of Psychonomics, University of Amsterdam, Roetersstraat 15, NL-1018 WB Amsterdam, The Netherlands. E-mail: m@meeter.nl;jaap@murre.com

The authors wish to thank Kim Graham for many stimulating discussions and her helpful comments while preparing this manuscript.

http://www.tandf.co.uk/journals/pp/09658211.html DOI:10.1080/09658210244000658

Patterson, Oxbury, & Funnell, 1992) and the temporal variant of frontotemporal dementia (Edwards Lee et al., 1997). Semantic dementia is associated with non-Alzheimer degenerative pathology of especially the inferolateral temporal neocortex, with relative sparing (at least in the early stages) of medial temporal regions (Hodges, Garrard, & Patterson, 1998; Mummery, Patterson, Price, Ashburner, Frackowiak, & Hodges, 2000); though some damage in the left hippocampus is evident in certain patients (Galton et al., 2001). Patients typically show a progressive deterioration in their semantic knowledge, and yet seem to possess relatively preserved day-to-day (episodic) memory. Therefore, semantic dementia provides a unique opportunity to investigate the organisation of human long-term memory.

Semantic dementia is primarily a disease in which semantic memory is compromised. Patients are unable to name previously familiar objects, people, and places, and show poor language comprehension. They show deficits on verbally based semantic memory tests such as category fluency and picture naming, but also on non-verbal tests of semantic memory: they have difficulty matching animal and object sounds to pictures of the animal or object (Bozeat, Lambdon Ralph, Graham, Patterson, Wilkin, & Rowland, 2000), and may have difficulties handling previously familiar objects (Hodges et al., 1998).

Episodic memory is relatively spared in semantic dementia. Nevertheless, it is not normal. Murre, Graham, and Hodges (2001) have recently discussed semantic dementia in the context of neuroanatomically based computational models of long-term memory, and suggested several characteristics of semantic dementia that these models could be expected to simulate appropriately. Three of the particularly salient ones were related to episodic memory.

1. Relative sparing of recent versus remote memories

Patients with semantic dementia show some retrograde amnesia. The amnesia is far more pronounced for the distant past than for recent periods. This has been demonstrated in both the autobiographical (Graham & Hodges, 1997; Murre et al., 2001; Snowden, Griffiths, & Neary, 1996) and public knowledge domains (Hodges & Graham, 1998; Snowden et al., 1996). These results stand in contrast to the Ribot gradient typically found in patients with an amnesic syndrome, in which remote memories are preferentially affected and recent memories are relatively spared.

Direct comparisons of semantic dementia with amnesia have generally been restricted to patients with Alzheimer's disease. Though Alzheimer's patients show pathology in almost all brain regions, the hippocampus is disproportionately affected in the early stages (Braak & Braak, 1991),

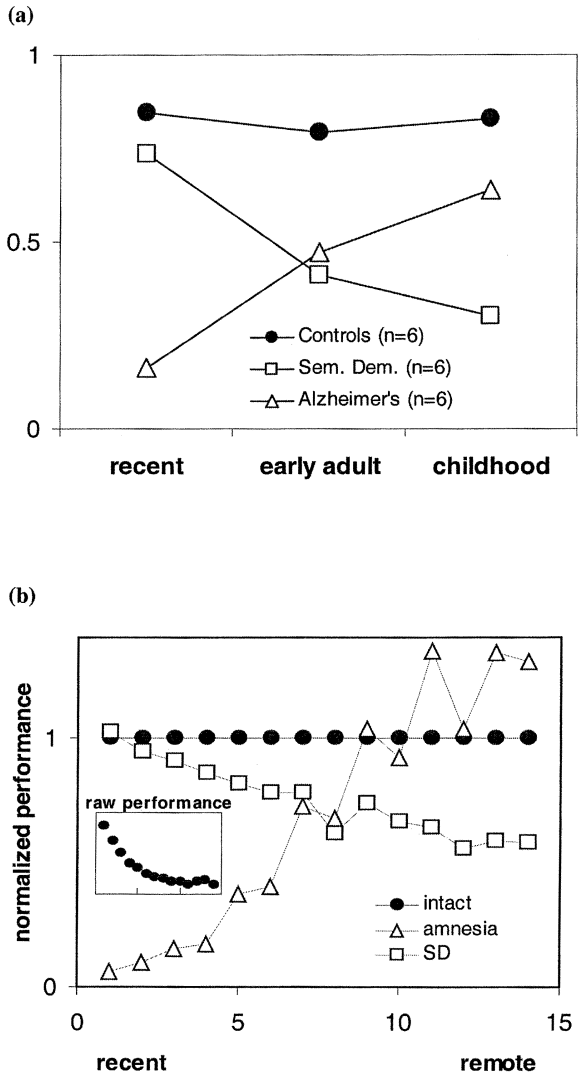


Figure 1. (a) Remote memory gradients from a study in which patients with Alzheimer's (AD), semantic dementia patients (SD), and controls (Ctrl) performed a remote memory task, the AMI-autobiographical incidents (adapted from Graham & Hodges, 1997). (b) Results of the remote memory gradient simulation. The x-axis gives the pattern number, listed from the most recent to the left, to the first-learned to the right. Performance in the amnesia and the semantic dementia (SD) conditions is presented as a fraction of the performance in the control condition with an intact model, which is set to 1.

and mild to moderate Alzheimer's patients have many characteristics in common with patients with discrete hippocampal damage (Deweer, Pillon, Pochon, & Dubois, 2001). Figure 1a shows one such comparison of a group of semantic dementia patients with normal controls and patients with Alzheimer's disease on a test of remote autobiographical memory (Graham & Hodges, 1997). While the Alzheimer's patients, like other amnesic groups (Gade & Morteson, 1990; Kopelman, 1989; Reed & Squire, 1998), were especially impaired for the most recent period, semantic dementia patients showed relatively preserved recent memory, and were impaired on remote memories.

2. Preservation of new learning, as measured by recognition memory, early in the disease

One of the consistent findings in semantic dementia is that patients, even those with severe semantic memory deficits, are still able to acquire new episodic memories. By contrast, an episodic learning deficit (anterograde amnesia) is the most prominent symptom of amnesic disorders. One semantic dementia patient, for example, who had forgotten the names of vegetables, was able to relearn to name them with the help of short descriptions (Funnell, 1995). Group studies show that patients with semantic dementia can perform normally on a three-alternatives forced-choice test of visual recognition memory (Graham, Becker, & Hodges, 1997; Graham, Patterson, & Hodges, 1999a; Graham, Simons, Pratt, Patterson, & Hodges, 2000, see Figure 2a). Although on the Rey complex figure their performance was not normal, it was far less impaired than that of a group of patients with presumed early Alzheimer disease (Graham et al., 2000). This capacity for new learning becomes compromised, however, as the disease progresses (Graham et al., 1999a).

3. Increased long-term forgetting of newly learned material

Murre (1996, 1997) has predicted, on the basis of an analytical review of the TraceLink model, that patients with semantic dementia might experience increased forgetting. Some evidence for this prediction came both from a recent semantic dementia case study (Graham, Patterson, Pratt, & Hodges, 1999b), and from cases of lateral temporal lobe lesions unrelated to semantic dementia (Kapur, Scholey, Moore, Barker, Brice, & Thompson, 1996). In two case studies (Graham et al., 1999b; Graham, Patterson, Pratt, & Hodges, 2001), a patient was able to relearn, via repeated training sessions, forgotten exemplars for categories, so as to perform normally on category fluency tests. Once the training sessions ceased, however, the exemplars were quickly forgotten. This is the opposite pattern of that found in amnesic patients, who once they have acquired memories, tend to show forgetting similar to that of normal controls (Huppert & Piercy, 1978; Kopelman, 1985).

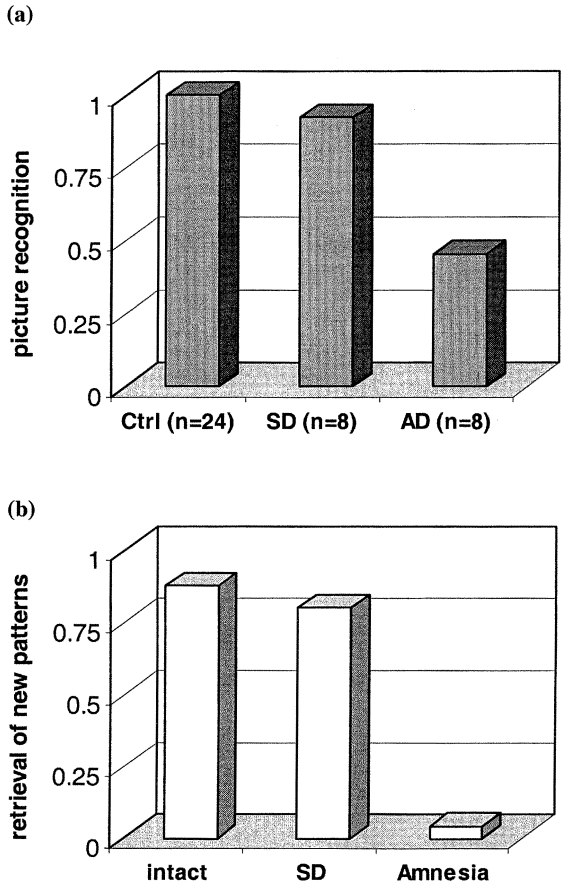


Figure 2. (a) Episodic learning in normal controls (Ctrl), patients with semantic dementia (SD), and patients with probable early Alzheimer's disease (AD). Plotted is performance, corrected for guessing, on three-alternative forced-choice recognition of 40 pictures (adapted from Graham et al., 2001). (b) Raw performance (porportion of recalled pattern nodes) on patterns 9, 10, and 11 immediately after acquisition. In the semantic dementia (SD) and amnesia conditions, these are the patterns acquired right after the lesion. ''Intact'' refers to the control simulation with the intact model.

AIM OF THE CURRENT RESEARCH

One can conclude that in the realm of episodic memory, semantic dementia has many characteristics opposite to those seen in amnesia (Graham et al., 1999a; Hodges, 1995; Hodges et al., 1998; Hodges et al., 1992; Snowden, Goulding, & Neary, 1989). Semantic dementia patients remain capable of new learning, and their retrograde amnesia tends to spare recent memories. This syndrome, therefore, offers a challenge to computational models of human memory. Only if they are able to simulate the neuropsychological phenomena observed in semantic dementia as well as in amnesia, can they be regarded as viable models of long-term memory. Although Murre et al. (2001) suggest, on the basis of verbal arguments, that connectionist models of remote memory and amnesia would be able to account for the data reviewed above, this has so far not been shown. The aim of the current paper is therefore to investigate whether one computational model of amnesia, TraceLink (Murre, 1996; Murre & Meeter, 2002), can simulate the findings.

In the next section, the TraceLink model will be explained. This explanation will focus on the conceptual level (details of the connectionist model are discussed in an appendix). The simulations of semantic dementia will be discussed thereafter.

THE TRACELINK MODEL

Structure of the model

The TraceLink model is one of the computational models that have been used to simulate remote memory and amnesia. A schematic drawing of the model is shown in Figure 3. Its three main components are (1) the trace system, (2) the link system, and (3) the modulatory system.

Trace system. Ultimately, the greater part of a memory trace will be stored in the trace system. The trace system represents roughly the neocortical basis of memories. The input to the trace system originates in sensory areas, which themselves do not form part of the model. Similarly, output or motor areas are not included, but rather are assumed. The trace system is identified with association areas in the neocortex that have been associated with memory, such as the temporal lobe neocortex (Miyashita, 1993) and posterior parietal cortex (Izquierdo et al., 1997).

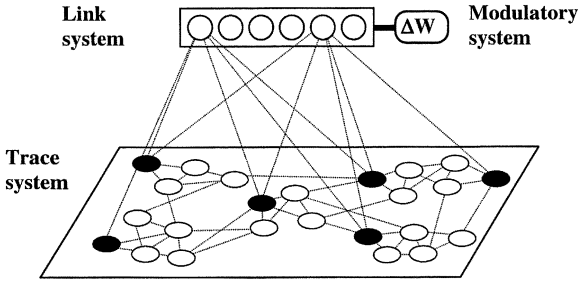


Figure 3. Overview of the TraceLink model, showing the link system, the trace system, and the modulatory system (indicated by a ΔW sign, symbolising control of learning rate on the connection weights in the system). Only a few nodes and a few connections have been drawn in order to prevent clutter in the drawing.

Link system. It is the link system's function to connect remote trace elements (i.e., those without direct cortico-cortical connections). The link system has a much smaller number of elements than the trace system. Among others, this implies that link elements are more likely to be reassigned from an old representation to a new, thus causing interference by new learning. Link elements are connected to each other and to a random subset of the trace elements. Connections involving link elements are much more plastic than trace-to-trace connections. This is illustrated in Figure 3 through the close attachment of the link system to the modulatory system. The link system can be identified with the hippocampus and other medial temporal lobe structures, such as the entorhinal cortex.

Modulatory system. Activation of the modulatory system causes increased plasticity of the link system. The modulatory system may be activated through central states such as arousal and attention, and through stimulus properties such as intensity or novelty. In the brain, several systems have been proposed that modulate learning, for example an amygdaloid system based on norepinephrine (McGaugh, 1990) and a cholinergic system arising from the basal forebrain (Hasselmo, 1995; for more modulatory systems see that paper). These are not implemented in this version of the model. The modulatory system can be identified with the basal forebrain and other subcortical centres that control plasticity in the brain.

Figure 4 illustrates at a conceptual level the formation of an episodic memory trace under normal circumstances. It is convenient to distinguish four stages:

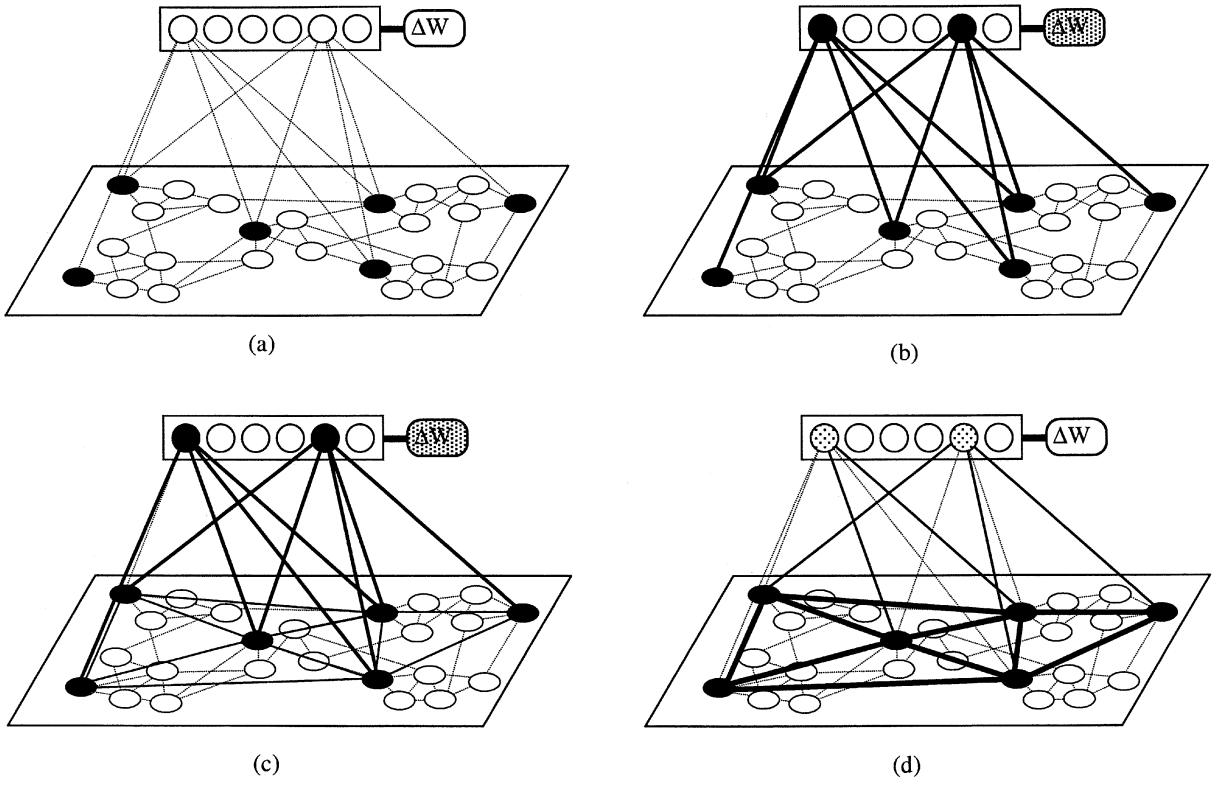


Figure 4. Four stages in the normal formation of episodic memories in the TraceLink model. *Stage 1:* A new memory representation activates a number of trace elements (shown as filled black circles). *Stage 2:* Several link elements are activated and the relevant trace–link connections are strengthened (shown as thicker connections). Also, the modulatory system has been activated. *Stage 3:* Weak trace–trace connections are developing. The modulatory system is weakly activated. *Stage 4:* Strong trace–trace connections are formed. Trace–link connections have decayed and the modulatory system does not necessarily respond to the stimulus.

Stage 1. Through the sensory and motor channels a set of trace nodes is activated (filled circles). This represents an episode to be remembered. The activated trace nodes have no direct trace-to-trace connections to each other but they are connected to a number of link nodes (only two are drawn). A mechanism of adaptable local inhibition keeps the average number of active nodes at some preset low level.

Stage 2. The trace nodes activate a set of link nodes. Also here, dynamic inhibition suppresses the link nodes with the weakest activation. The modulatory system becomes activated (darkness of shading indicates activation level) and the learning rate increases. As a result of the increased plasticity, connections between link and trace nodes are strengthened (shown by a thickening of the connections). This can take place in seconds or minutes.

Stage 3. Repeated activation of the memory trace through the link system will lead to the gradual formation of trace-to-trace connections. In the brain, this probably does not involve just the formation of long-distance synapses, but also and to a larger extent the formation of chains of connected neurons that bridge the gap between two distant areas (so-called *Abeles chains*; Abeles, 1991). This process is identified with consolidation. As the pattern is now old, the modulatory system will now no longer react strongly to the pattern.

Stage 4. Direct trace-to-trace connections have become very strong. Link–trace connections have either decayed or have been reassigned to other memory traces. The modulatory system shows little reaction to the pattern. In short, the memory trace has become independent of the link system.

Recall is modelled as the retrieval of the whole trace pattern when part of the trace pattern is offered as a cue. In the case of a stage 2 memory, the partial cue will typically activate the link nodes

associated with the pattern, which in turn will activate the rest of the trace pattern. In the case of a stage 4 memory, the cue will be able to activate the rest of the trace pattern directly through strong trace–trace connections.

Consolidation is conceptualised as an autonomous process. Although there are strong hints that consolidation may occur in sleep (Dave & Margoliash, 2000; Frank, Issa, & Stryker, 2001; Karni, Tanne, Rubinstein, Askenasy, & Sagi, 1994; Philal & Born, 1999; Smith, 1995; Stickgold, James, & Hobson, 2000; Van Ormer, 1932; Wilson & McNaughton, 1994), the TraceLink model is not tied to this assumption. Consolidation is implemented in the model as a competitive process consisting of discrete consolidation trials. In each trial, one pattern becomes activated by letting the model find an attractor from a random cue. The trace portion of this attractor (usually a stored pattern) is then strengthened. As recent memories are strong in the link system, these memories often win the competition for consolidation. While these memories gradually lose their link system representation, strength is built up in the trace system through consolidation, allowing memories to remain retrievable.

The structures that are part of the link system, mainly medial temporal lobe structures, are often implicated in amnesia (Bauer et al., 1993; Squire, 1992; Square & Zola-Morgan, 1991). Amnesia is thus modelled in TraceLink by a lesion to the link system. Such a lesion produces both anterograde amnesia and retrograde amnesia with a Ribot gradient. Why this is the case can be understood in terms of the stages illustrated in Figure 4. If the link system is lesioned, new patterns in the trace system cannot be bound together by link system nodes. These patterns will thus not be able to enter stage 2, the stage where they can be recalled via the link system. This explains the anterograde amnesia. Memories formed just before the lesion will still be in stage 2. Since stage 2 memories depend for their retrieval on the intact link system, these new memories will be lost after the lesion. Old memories—those in stage 4—will still be retrievable, however, as the connections within the trace system will enable these memories to be retrieved. Stage 3 memories will be somewhere in between, with some having strong enough connections within the trace system to be retrieved and others not. This explains the Ribot curve: old memories can be retrieved after a link lesion, recent memories cannot.

Other simulations have shown that the TraceLink model can explain, among other things, temporary amnesia such as transient global amnesia (TGA), shrinking of amnesia, the intactness of implicit memory in amnesia, isolated retrograde amnesia, the effects of arousal on consolidation, the shape of normal forgetting, and the existence of permastore (Murre & Meeter, 2002).

The model used in the simulations

In the simulations described in this paper, the same set-up and parameters are used as were used previously to simulate amnesia (Murre & Meeter, 2002). The simulation set-up will briefly be discussed here; for technical details, the reader is referred to the Appendix.

The trace system is modelled as a layer of 200 nodes, the link system as a layer of 42 nodes. Both layers have internal connections, and are connected with each other. Between every two nodes an excitatory connection can in principle be formed. Since none of the simulations depends on more than a coarse notion of modulation, we do not model the modulatory system of the TraceLink model explicitly, but instead we merely model its function.

Both layers have binary stochastic nodes which model groups of neurons that are at some distance from each other. The likelihood with which a node is active depends on the balance between the excitatory input from other nodes and inhibition. The excitatory input to a node is the weighted sum of the activation of all nodes connected to it. Inhibition is constantly fine-tuned so as to keep the average number of active cells in a layer as close to a preset number (k) as possible: inhibition is increased when too many nodes are active (i.e., more than k), and lowered when too few are active (i.e., less than k). The number k is set separately for every layer, and consequently inhibition is regulated separately in every layer. The weights of excitatory connections are modifiable with a variant of Hebb's rule (Hebb, 1949; Singer, 1990) that allows for learning as well as unlearning as a function of contingent and non-contingent activity of the nodes. Weights can vary between 0 and 1, and are changed during learning in a simple, linear fashion.

Learning is not equally fast for all connections. The learning rate—the rate at which changes are made to the weights during learning—is much

lower for the within-trace connections than for the connections within the link layer, or between the link layer and the trace layer; connections between the trace and link layers and within the link layer have equal learning rates. This models the fact that the hippocampus is unusually plastic (Lopes da Silva, Witter, Boeijinga, & Lohman, 1990), and that the regions of the link system have a higher connectivity than the neocortex (Treves & Rolls, 1994).

A pattern in the learning set consists of a random group of trace nodes and link nodes, which are activated when a pattern is presented. The number of nodes in the trace layer and the link layer that belong to a pattern is equal to the number of nodes active in the layer in equilibrium (k).

Simulating normal memory

In all simulations, the model goes through two distinct phases, a learning phase and a test phase (see Figure 5). The learning phase consists of two alternating sub-phases, acquisition of a pattern and consolidation. During acquisition, the model learns one new pattern. This is followed by a period of consolidation, after which another pattern is acquired.

In a consolidation period, three consolidation trials occur (see Figure 5).[1] In a single consolidation trial the model is allowed to cycle freely for a fixed number of iterations (150), and whichever pattern is active at the last iteration is consolidated. The stochastic dynamics of the model thus choose a pattern to consolidate. The model usually does not settle on the same pattern for all three trials in a consolidation period, and thus more than one pattern is consolidated in one consolidation period.

After learning and consolidation, the model is subjected to a test phase. Each pattern is tested a number of times by activating and clamping part of the pattern in the trace layer (a cue of three nodes), and then letting the model cycle for 70 iterations. No node is clamped in the link layer. The proportion of pattern nodes not part of the cue that are active after the 70th iteration is used as the measure of performance.

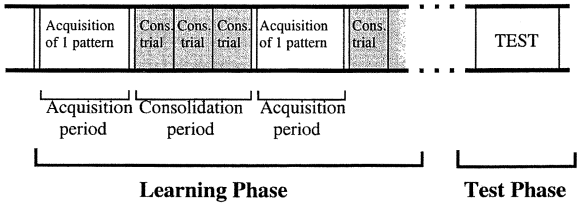


Figure 5. Diagram showing the order of events in most simulations. A simulation was divided into a learning phase and a test phase. The learning phase was subdivided into alternating acquisition periods, in which one pattern was acquired, and consolidation periods. Consolidation periods consisted of three consolidation trials.

[1] To ensure that the first patterns do not get too large a head start, there are fewer than three consolidation trials after acquisition of the first two patterns. As there are no previous patterns for the model to choose, the first patterns are always consolidated in the first consolidation trial. Since the very first pattern is learned in an ''empty brain'', it is always excluded from any analysis or figure showing results.

Simulating amnesia and semantic dementia

As already stated, amnesia can be modelled in the TraceLink model by a loss of link nodes. This has different effects for new and recent memories, as compared to older ones. New memories can no longer be encoded in a stable form, because there are no link nodes that can bind together the trace nodes in the pattern. Recent memories that are still dependent on a functioning link system for their retrieval (stage 2) will be lost, while those memories that have developed a supporting trace–trace connectivity structure (stage 4), can still be retrieved through those connections.

Semantic dementia can be simulated as a loss of trace elements and of trace–trace connections (i.e., cortico-cortical connections). This mimics the atrophy of temporal neocortex seen in semantic dementia (Galton et al., 2001; Hodges et al., 1998; Mummery et al., 2000). At the onset of the disease, loss of trace–trace connections dominates, as the chains connecting distant regions of the cortex (Abeles chains) will be more vulnerable than the functional units in the neocortex (i.e., hypercolumns or ensembles). In Figure 6, the formation of memory traces is drawn in case there is severe damage to the trace–trace connections in the trace system, with sparing of a majority of trace elements (Figure 6a). Stage 1 memories can be transformed into stage 2 memories (figure 6b), as the link system and modulatory system are still fully operational. The transition to stages 3 and 4, however, is severely impaired, because there are not enough trace–trace connections left to form supporting networks at the trace level (Figure 6c). The system

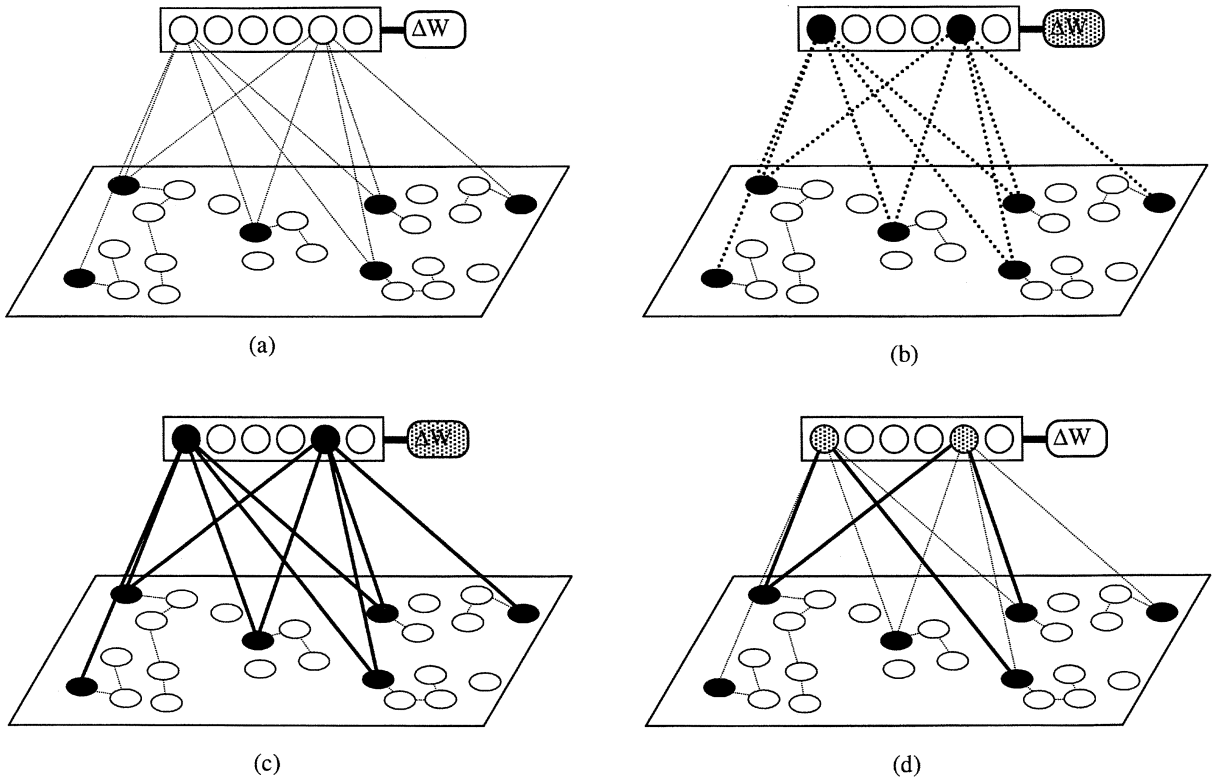


Figure 6. The four stages of memory in the TraceLink model in semantic dementia. See text for explanation.

will show (*i*) a diffuse but possibly extensive loss of existing well-consolidated memories, (*ii*) preservation of the formation of episodic memories through the link system, (*iii*) strong interference of new over old episodic memories, because of the limited capacity of the link system (figure 6d). This will cause learned episodes to be forgotten faster than in the intact model. If a memory is not rehearsed regularly it will be lost from the link system. This behaviour is very similar to that observed in patients with semantic dementia.

In the next section, these explanations will be further explored in a series of simulations.

THE SIMULATIONS

Three simulations of semantic dementia were carried out, addressing the three properties of episodic memory in semantic dementia discussed above:

- relative sparing of recent versus remote memories;
- intact new learning;
- accelerated forgetting of newly learned material.

In each of these simulations, the behaviour of the intact model was compared to an ''amnesic'' model (with link deactivation) and to a ''semantic dementia'' model in which lesions to trace were made.

Simulation 1: Remote memory gradient

Method. In the first simulation, the retention curve was tracked under three conditions: normal memory, retrograde amnesia, and semantic dementia. In all three conditions, a list of 15 patterns was learned by the model, with consolidation periods interspersed between pattern acquisitions. Then the model was tested three times (since no learning occurred during the tests, preceding tests do not alter the results of subsequent tests; the three tests are therefore independent, and can be considered as within-subject comparisons). It was first tested in intact state for the control condition. Then, the link layer was deactivated, and the model was tested again for the amnesia condition.

To simulate a state of progressed semantic dementia, trace nodes and connections within the trace layer were lesioned. Our standard lesion was

80% of all connections within the trace layer and 10% of trace nodes (in simulation 3 other lesion sizes will be explored). The model was tested a third and final time in this state in the semantic dementia condition. This completed the simulation.

Aside from the new testing condition (the semantic dementia condition), no changes were made compared with previous simulations of amnesia (Murre & Meeter, 2002). No assumptions were added, nor were there any new parameters besides those involved in the semantic dementia manipulation.

Results and discussion. In the control condition, the model shows the power law forgetting curve found in many studies of human memory (Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991; see Figure 1b, inset).

Constructors of retrograde amnesia tests usually attempt to hold performance constant across decades for normal controls (Hodges, 1995; Mayes, Downes, McDonald, Rooke, Sagar, & Mendell, 1994). As a consequence of this, memories for different time periods are not equivalent in original learning strength. The inequality of the items for different time periods can best be simulated by calculating the performance of the abnormal groups as a fraction of the performance of the normal controls. For better comparison with data using retrograde amnesia tests, we therefore report the results of the amnesia and semantic dementia as a fraction of the performance in the control condition.

As can be seen in Figure 1b, deactivating the link layer in the retrograde amnesia condition generates a Ribot curve: recent patterns are lost, but remote ones were still available after deactivation of the link layer. In the semantic dementia condition, however, exactly the opposite occurs: the most recent patterns are still available, while old patterns are severely degraded by the lesion of the trace–trace connections. This amounts to a reverse pattern of retrograde amnesia similar to the one observed in semantic dementia patients (see Figure 1a for comparison).

Simulation 2: New leanring and forgetting

Method. To test whether learning would remain intact in a state of simulated semantic dementia, we compared learning in the intact model with learning in a condition where a lesion to trace system connectivity had been made. In the semantic dementia condition, a lesion of again 80% of all trace–trace connections and 10% of trace nodes was made after the acquisition of eight patterns. Subsequently, eight more patterns were presented to the model in its lesioned state. In the control condition, all 16 patterns were learned by an intact model. In both simulations, the model learned the patterns by going through the alternation of acquisition and consolidation as explained above.

Results and discussion. We analysed both new learning, and the forgetting of patterns in the model. To measure new learning, performance on newly acquired patterns was tested immediately after acquisition. In Figure 2b, average retrieval of the first three patterns acquired after the lesion is plotted for the three conditions. The semantic dementia lesion did not hamper the ability of the model to encode new patterns, although new learning was much diminished in the amnesia condition.

To investigate forgetting in the model, we plotted how well the patterns acquired after the lesion (patterns 9, 10, 11 etc.) could be recalled after acquisition of more patterns, Figure 7 plots, as an example, how well pattern 10 can be retrieved after acquisition of the pattern itself, and after one pattern more is acquired, two patterns more, etc. To test whether forgetting was faster after a lesion of trace connectivity, we fitted power functions

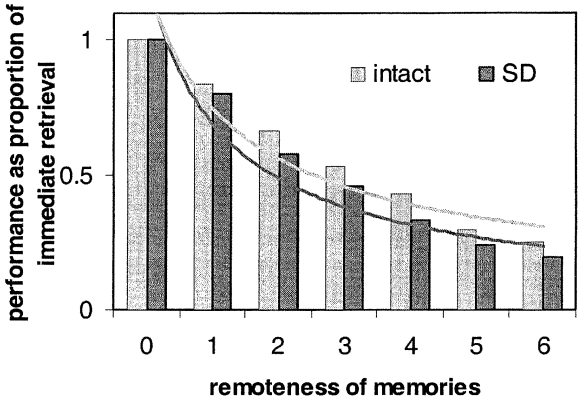
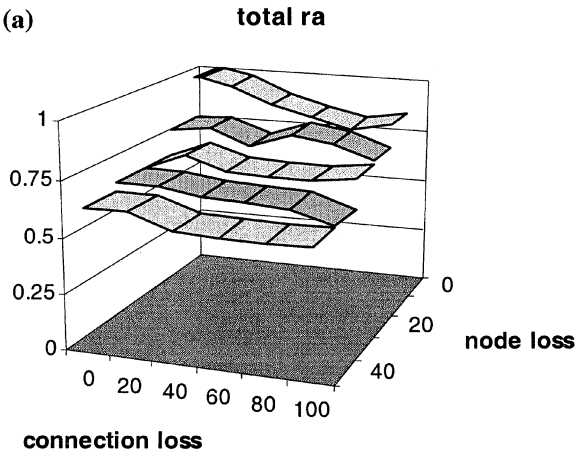


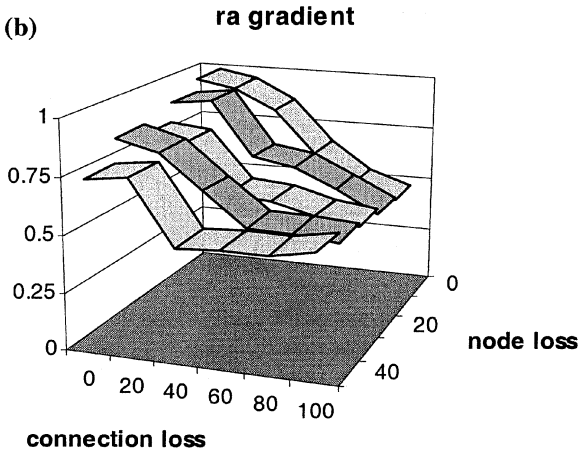
Figure 7. The retention curve of pattern 10 in both the control condition and the semantic dementia conditions. For both curves, performance immediately after acquisition is set to 1, and then retention of pattern 10 is measured after acquisition of subsequent patterns. The x-axis gives the number of interfering, newly acquired patterns. For example, at "1" the performance on pattern 10 is given after pattern 11 has been acquired.

through all observed forgetting curves (i.e., those for patterns 9 to 15). Then we compared the exponent, which determines the rate of decay of a power forgetting function, of the best fitting power function for the forgetting curves of all patterns. For all seven patterns, the exponent was larger—and therefore forgetting faster—in the semantic dementia condition than in the control condition: the exponent for the control curves ranged from -0.36 to $-.039$, while it ranged from -0.42 to -0.45 for the semantic dementia curves ($p < .01$ with a simple binomial test of significance).

The simulation shows that the model was still capable of storing new patterns in the semantic dementia condition. Although acquisition was not worse than in the control condition, patterns were lost more rapidly in the semantic dementia condition. This is analogous to what is found in patients with semantic dementia, where intact episodic learning is combined with speeded forgetting (Graham et al., 1999b).

In the TraceLink model, these results follow naturally from the assumption that there is a fast-learning temporary repository (the link system) from where patterns are consolidated into a more permanent store (the trace system). After the lesion of trace system connectivity with which we model temporal neocortex atrophy lesions, the model is still able to store patterns through the link system. Episodic learning is therefore intact. However, patterns are lost fairly rapidly in the link system, and as patterns can only be consolidated in a rudimentary form if connections within the trace system are lesioned, patterns cannot build up a more permanent representation in that system.





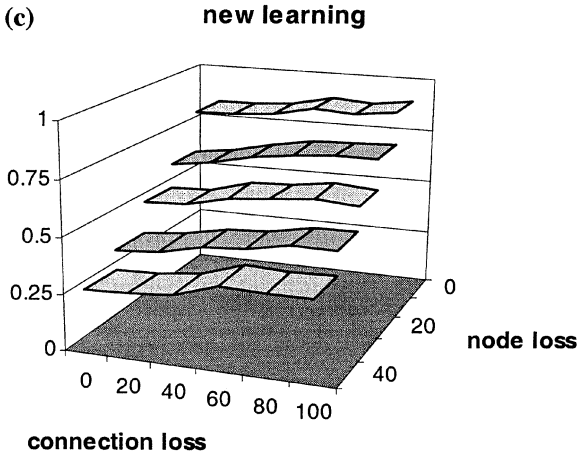


Figure 8. (a) Retrograde amnesia in the semantic dementia condition for varying lesion sizes (performance on stored patterns in the semantic dementia condition as a proportion of performance in the control condition). (b) Gradient of retrograde amnesia in semantic dementia, measured as retrograde amnesia for remote patterns (patterns 1 to 4) divided by retrograde amnesia for recent patterns (patterns 11 to 14). For ungraded amnesia this number is 1, for amnesia with a reverse gradient it is lower than 1. The lower the number, the steeped the reverse gradient. (c) New learning, as measured by raw average performance on the first three patterns learned after the lesion (same as in Figure 2).

Simulation 3: Varying the lesion size

In the previous simulations, the semantic dementia condition consisted of a lesion of 80% of trace–trace connections, and 10% of trace nodes. To investigate whether this particular choice of lesion influenced the results, we redid simulations 1 and 2 with a large variation in lesion sizes. This also offers an opportunity to investigate how TraceLink would predict memory deficits to progress with progression of temporal lobe lesions.

Method. The simulations 1 (remote memory) and 2 (new learning) were repeated, with the semantic dementia lesion varying in size. The percentage of connectionns lesioned varied between 0 and 100%, the number of trace nodes lesioned between 0 and 40%.

Results and discussion. Figure 8 shows how retrograde amnesia, the gradient in retrograde amnesia, and new episodic learning depend on the size of the trace lesion. As our measure of retrograde amnesia, we summed performance on all patterns stored before the lesion and divided them by the same performance in the control condition (Figure 8a). We then divided relative performance on the most recent patterns, by relative performance on the remote memories (both were first divided by performance in the control simulation). This gave us an index of the gradient: a value of 1 indicates that all memories were equally affected, while lower values indicate a reverse gradient (Figure 8b). Our measure of new learning, finally, was the same as that in the previous simulation: performance on the first three patterns learned after the lesion (Figure 8c).

Retrograde amnesia (Figure 8a) became progressively worse with increasing lesions of trace nodes and trace–trace connections. Its gradient also became steeper with more lesioned trace nodes, and especially with more lost connections (Figure 8b). New learning, on the other hand, held up very well to lesions of connections, but suffered strongly from lesions of trace nodes (Figure 8c).

One way to interpret these results is that broad lesions (loss of connectivity) are responsible for the reverse gradient in remote memory, while localised lesions (loss of nodes) cause a loss of new learning capacity. If progression of the disease can be seen as a slow movement from the upper left corner in the plots of Figure 8 (no lesion) to the lower right corner (many nodes and connections lost), then the present results would predict that the gradient in remote memory comes early. With larger lesions, the steepness of the reverse gradient reaches a plateau, and remote and recent memories degrade at an equal rate. The capacity for new learning is initially spared, but becomes compromised with more sustained lesions, in which trace nodes are lost. Indeed, episodic learning becomes compromised later in the disease (Graham et al., 1999a).

GENERAL DISCUSSION

The TraceLink model was originally developed as a model for amnesic syndromes (Murre, 1996, 1997; Murre & Meeter, 2002). Here we demonstrate that the model, without any additional assumptions or additional parameters, can reproduce some of the principal characteristics of episodic memory in semantic dementia. A lesion of trace nodes and trace–trace connections, modelling atrophy of primarily temporal neocortex, was shown to produce:

- a relative sparing of more recent memories in retrograde amnesia;
- intact new learning of patterns;
- speeded forgetting of newly acquired memories.

From our third simulation, some predictions can be generated about how the disease progresses. Our results predict that a graded loss of remote memories comes earlier in the disease (when most loss involves connectivity), with a broad retrograde amnesia and a loss of learning capacity following with more extensive damage.

In a fourth, explorative simulation not reported here, we investigated implicit learning. We modelled it as learning involving only the trace system, in accordance with a large body of theoretical ideas and empirical data implicating neocortical processing areas in implicit memory, and implicit memory's independence of medial temporal lobe regions (Gabrieli, 1998; Keane, Gabrieli, Mapstone, Johnson, & Corkin, 1995; Knowlton & Squire, 1993). An implicit learning trial resulted in improved performance for the control and amnesia conditions. In the semantic dementia condition, performance also improved after an implicit memory trial, but markedly less so than in the other conditions. TraceLink thus predicts a priming deficit in semantic dementia for those types of implicit learning that involve the lateral temporal lobe. Although perceptual priming seems to be intact in semantic dementia (Srinavas, Breedin, Coslett, & Saffran, 1997), impaired conceptual priming was indeed found in patients with semantic dementia (Moss, Tyler, Hodges, & Patterson, 1995; Nakamura, Nakanishi, Hamanaka, Nakaaki, & Yoshida, 2000). However, the semantic priming measures used in those studies were different from the repetition priming simulated in TraceLink.

All findings were simulated on a qualitative rather than qualitative level. However, no new assumptions needed to be made to extend the model from amnesia to semantic dementia, and no parameter needed tuning (the direction of the effects was independent of the size of the semantic dementia lesion). Other models of amnesia, notably those of Alvarez and Squire (1994) and McClelland et al. (1995), would probably also be able to simulate some—if not all— of the findings. This is so because the main assumption that enables TraceLink to simulate them, namely consolidation of memories to a neocortical memory store, is shared by all three models. However, the Alvarez and Squire model is probably too small to produce the effects on gradients and of new learning (as only two patterns can be stored in the model), and the McClelland et al. model does not implement the hippocampal memory system (see Murre et al., 2001, for a more thorough discussion of the similarities and differences between these models).

A limitation of the current model is that it does not incorporate the more basic, visual systems that probably play a role in intact episodic learning, and intact perceptual priming. Primary and secondary sensory cortices seem to be the locus of perceptual priming effects (Cabeza & Nyberg, 2000; Gabrieli, 1998; Keane et al., 1995), while episodic encoding and retrieval of visual stimuli often causes activation in occipital, visual cortex (Cabeza & Nyberg, 2000). Primary and secondary sensory cortices have also been suggested to play an important role in intact learning in semantic dementia. Graham et al. (2000) state that intact learning seems primarily perceptual in nature, as recognition performance of semantic dementia patients in their experiment was only normal with perceptually identical stimuli. It was not normal when at test subjects were confronted with perceptually different stimuli that portrayed the same object as during study. Therefore, Graham et al. suggested, direct connections from these structures to the hippocampal memory system, bypassing the semantic system in the temporal lobe, might be responsible for intact learning of visual stimuli.

However, perceptual learning may be better than nonperceptual learning not because of direct connections between perceptual areas and the hippocampal system, but because a damaged temporal lobe may be able to sustain storage and retrieval longer for materials for which it is just the gateway than for materials for which it is the ultimate store (such a vocabulary). In this context, it is important to note that even with perceptually different stimuli, performance for semantic dementia patients was much better than that of Alzheimer's patients in Graham et al.'s study, and that a direct comparison of visual and verbal learning in semantic dementia patients has never been done. It is thus too early to decide whether intact episodic learning in semantic dementia requires new connections. Either way, inclusion of sensory areas that remain intact in semantic dementia in TraceLink would probably increase its explanatory power, for example by offering a way to simulate different forms of priming, and account for their dissociation in semantic dementia.

Although the model is thus far from complete, it can be concluded that recent findings from semantic dementia offer support for the view that memory consolidation in humans is dependent upon interactions between the hippocampal complex and the neocortex (Alvarez & Squire, 1994; Marr, 1971; McClelland et al., 1995; Murre & Meeter, 2002; Squire & Alvarez, 1995). This idea, although by now so common as to be called the ''Standard View'' of retrograde amnesia (Nadel & Moscovitch, 1997), is still controversial. Nadel, Moscovitch and co-workers (Nadel & Moscovitch, 1997; Nadel, Samsonovitch, Ryan, & Moscovitch, 2000) have developed a competing view, the multiple-trace theory, which assumes that the hippocampus is always necessary for episodic memory retrieval. After complete lesions of the hippocampus, Nadel and Moscovitch (1997) claim, the typical pattern is not a Ribot gradient, but ungraded retrograde amnesia spanning the whole lifetime of the patient. Temporal gradients, when they occur, are explained as stemming from partial lesions, which are more likely to lead to the loss of recent than remote memories in their model.

With regard to semantic dementia, some discussion has also ensued as to whether it genuinely offers support for the ''Standard View'' or not (Graham et al., 1999b; Moscovitch & Nadel, 1999). Moscovitch and Nadel (1999) have claimed that it does not, on the grounds that the data pattern is different from what is commonly assumed. Unlike TraceLink, the multiple-trace theory would predict that semantic dementia patients do not have a different remote memory curve from normal controls: although the performance on retrograde amnesia tests may be lower for patients, remote memories should not be

affected any more than recent memories. Indeed, Moscovitch and Nadel claim that this is the case. The gradients reported in the literature could be the result of a greater difficulty of old versus new items on the tests, which is masked in the performance of normal controls by ceiling effects (Nadel et al., 2000). Moreover, retrograde amnesia may be overestimated because of a reliance on verbal cueing. Visual cueing of memories might, they claim, lead to better recall of remote memories. This they support with observations from one patient with semantic dementia, who, with help of visual cues, could be made to remember his wartime memories of more than half a century ago (Nadel et al., 2000).

However, controls do not perform close to ceiling in all tests of remote memory. Hodges and Graham (1998) found evidence for a better performance of semantic dementia patients for recent versus remote memories on the identification of famous names, even though normal controls performed off the ceiling and were not better for recent than for remote periods. Although an assessment of remote memory with visual cues has yet to take place, we would argue that the evidence so far has been on the side of the reality of the reverse pattern in retrograde amnesia. As the pattern in the retrograde amnesia seen in semantic dementia can clearly differentiate between the Standard View espoused here and the multiple-trace theory, more research is certainly warranted.

If a reverse pattern is indeed present, it supports the assumptions underlying TraceLink. Together with the finding from semantic dementia, the simulations presented here makes a stronger case for the view that memories are, with help of the hippocampal memory system, consolidated into the neocortex, and that temporal neocortex atrophy blocks this consolidation process.

Manuscript received 6 March 2002
Manuscript accepted 1 November 2002
PrEview proof published online 25 July 2003

REFERENCES

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Alvarez, R., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of National Academy of Sciences (USA)*, *91*, 7041–7045.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396–408.

Bauer, R. M., Tobias, B., & Valenstein, E. (1993). Amnesic disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (3rd ed., pp. 532–602). New York: Oxford University Press.

Bozeat, S., Lambdon Ralph, M. A., Graham, K. S., Patterson, K., Wilkin, H., & Rowland, J. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*, 1207–1215.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathalogica*, *82*, 239–259.

Cabeza, R., & Ryberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.

Dave, A. S., & Margoliash, D. (2000). Song replay during sleep and computational rules for sensorimotor vocal learning. *Science*, *290*, 812–816.

Deweer, B., Pillon, B., Pochon, J. B., & Dubois, B. (2001). Is the HM story only a ''remote memory''? Some facts about the hippocampus and memory in humans. *Behavioural Brain Research*, *127*, 209–224.

Edwards Lee, T., Miller, B. L., Benson, D. F., Cummings, J. L., Russell, G. L., Boone, K. et al. (1997). The temporal varient of frontotemporal dementia. *Brain*, *120*, 1027–1040.

Frank, M. G., Issa, N. P., & Stryker, M. P. (2001). Sleep enhances plasticity in the developing visual cortex. *Neuron*, *30*, 1–20.

Funnell, E. (1995). A case of forgotten knowledge. In R. Campbell & M. Conway (Eds.), *Broken memories* (pp. 225–236). Oxford: Blackwell.

Gabrieli, J. D. E. (1998). cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*, 87–115.

Gade, A., & Morteson, L. (1990). Temporal gradient in the remote memory impairment of amnesic patients with lesions in the basal forebrain. *Neuropsychologia*, *28*, 985–1001.

Galton, C. J., Gomez-Anson, B., Antoun, N., Scheltens, P., Patterson, K., Graves, M. et al. (2001). Temporal lobe rating scale: Application to Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, *70*, 165–173.

Graham, K. S., Becker, J. T., & Hodges, J. R. (1997). On the relationship between knowledge and memory for pictures: Evidence from the study of patients with semantic dementia and Alzheimer's disease. *Journal of the International Neuropsychological Society*, *3*, 534–544.

Graham, K. S., & Hodges, J. R. (1997). Differentiating the roles of the hippocampal complex and the neocortex in long-term memory storage: Evidence from the study of semantic dementia and Alzheimer's disease. *Neuropsychology*, *11*, 77–89.

Graham, K. S., Patterson, K., & Hodges, J. R. (1999a). Episodic memory: New insights from the study of semantic dementia. *Current Opinion in Neurobiology*, *9*, 245–250.

Graham, K. S., Patterson, K., Pratt, K. H., & Hodges, J. R. (1999b). Relearning and subsequent forgetting

of semantic category examplars in a case of semantic dementia. *Neuropsychology*, *13*, 359–380.

Graham, K. S., Patterson, K., Pratt, K. H., & Hodges, J. R. (2001). Can repeated exposure to ''forgotten'' vocabulary help alleviate word-finding difficulties in semantic dementia? An Illustrative case study. *Neuropsychological Rehabilitation*, *11*, 429–454.

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia*, *38*, 313–324.

Hasselmo, M. E. (1995). Neuromodulation and cortical function: Modeling the physiological basis of behavior. *Behavioural Brain Research*, *67*, 1–27.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hodges, J. R. (1995). Retrograde amnesia. In A. D. Baddeley, B. A. Wilson, F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 81–107). Chichester, UK: Wiley & Sons.

Hodges, J. R., Garrard, P., & Patterson, K. (1998). Semantic dementia. In A. Kertesz & D. G. Munoz (Eds.), *Pick's disease and pick complex* (pp. 83–104). San Diego, CA: Wiley-Liss.

Hodges, J. R., & Graham, K. S. (1998). A reversal of the temporal gradient for famous person knowledge in semantic dementia: Implications for the neural organisation of long-term memory. *Neuropsychologia*, *36*, 803–825.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*, 1783–1806.

Huppert, F. A., & Piercy, M. (1978). Dissociation between learning and remembering in organic amnesia. *Nature*, *275*, 317–318.

Izquierdo, I., Quillfeldt, J. A., Zanatti, M. S., Quevedo, J., Schaeffer, E., Schmitz, P. K. et al. (1997). Sequential role of hippocampus and amygdala, entorhinal cortex and parietal cortex in formation and retrieval of memory for inhibitory avoidance in rats. *European Journal of Neuroscience*, *9*, 786–793.

Kapur, N. (1999). Syndromes of retrograde amnesia: A conceptual and empirical synthesis. *Psychological Bulletin*, *125*, 800–825.

Kapur, N., Scholey, K., Moore, E., Barker, S., Brice, J., & Thompson, S. (1996). Long-term retention deficits in two cases of disproportionate retrograde amnesia. *Journal of Cognitive Neuroscience*, *8*, 416–436.

Karni, A., Tanne, D., Rubinstein, B. S., Askenasy, J. J., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science*, *265*, 679–682.

Keane, M. M., Gabrieli, J. D. E., Mapstone, H. C., Johnson, K. A., & Corkin, S. (1995). Double dissociation of memory capacities after bilateral occipital lobe or medial temporal lobe lesions. *Brain*, *118*, 1129–1114.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, *262*, 1747–1749.

Kopelman, M. D. (1985). Rates of forgetting in Alzheimer-type dementia and Korsakoff's syndrome. *Neuropsychologia*, *15*, 527–541.

Kopelman, M. D. (1989). Remote and autobiographical memory, temporal context memory, and frontal atrophy in Korsakoff and Alzheimer patients. *Neuropsychologia*, *27*, 437–460.

Lopes da Silva, F. H., Witter, M. P., Boeijinga, P. H., & Lohman, A. H. (1990). Anatomic organization and physiology of the limbic cortex. *Physiological Reviews* , *70*, 453–511.

Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society B*, *262*, 23–81.

Mayes, A. R., Downes, J. J., McDonald, C., Rooke, S., Sagar, H. J., & Meudell, P. R. (1994). Two tests for assessing remote public knowledge: A tool for assessing retrograde amnesia. *Memory*, *2*, 183–210.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McGaugh, J. L. (1990). Significance and remembrance: The role of neuromodulatory systems. *Psychological Science*, *1*, 15–25.

Milner, P. M. (1989). A cell assembly theory of hippocampal amnesia. *Neuropsychologia*, *6*, 215–234.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of the Neurosciences*, *16*, 245–263.

Moscovitch, M., & Nadel, L. (1999). Multiple-trace theory and semantic dementia: Response to K. S. Graham (1999). *Trends in the Cognitive Sciences*, *3*, 87–89.

Moss, H. E., Tyler, L. K., Hodges, J. R., & Patterson, K. (1995). Exploring the loss of semantic memory in semantic dementia: Evidence from a primed monitoring study. *Neuropsychology*, *9*, 16–26.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*, 36–45.

Murre, J. M. J. (1996). TraceLink: A model of amnesia and consolidation of memory. *Hippocampus*, *6*, 675–684.

Murre, J. M. J. (1997). Implicit and explicit memory in amnesia: Some predictions by the TraceLink model. *Memory*, *5*, 55–82.

Murre, J. M. J. (2002). Connectionist models of memory disorders. In A. D. Baddeley, B. Wilson, & M. Kopelman (Eds.), *Handbook of memory disorders* (2nd ed., pp. 101–122). New York: John Wiley.

Murre, J. M. J., Graham, K. S., & Hodges, J. R. (2001). Semantic dementia: Relevance to connectionist models of long-term memory. *Brain*, *124*, 647–675.

Murre, J. M. J., & Meeter, M. (2002). *TraceLink: A connectionist model of consolidation and amnesia*. NeuroMod Technical Report 02-03, neuroMod group, Dept. of Psychology, University of Amsterdam.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217–227.

Nadel, L., Samsonovitch, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging and neuropsychological results. *Hippocampus*, *10*, 352–368.

Nakamura, H., Nakanishi, M., Hamanaka, T., Nakaaki, S., & Yoshida, S. (2000). Semantic priming in patients with Alzheimer and semantic dementia. *Cortex*, *36*, 151–162.

Phital, W., & Born, J. (1999). Memory consolidation in human sleep depends on inhibition of glucocorticoid release. *Neuroreport*, *10*, 2741–2747.

Reed, J. M., & Squire, L. R. (1998). Retrograde amnesia for facts and events: Findings from four new cases. *Journal of Neuroscience*, *18*, 3943–3954.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734–760.

Singer, W. (1990). Ontogenetic self-organization and learning. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain organization and memory: Cells, systems, and circuits* (pp. 211–233). Oxford: Oxford University Press.

Smith, C. (1995). Sleep states and memory processes. *Behavioural Brain Research*, *69*, 137–145.

Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioral Neurology*, *2*, 167–182.

Snowden, J. S., Griffiths, H., & Neary, D. (1996). Semantic–episodic memory interactions in semantic dementia: Implications for retrograde memory function. *Cognitive Neuropsychology*, *13*, 1101–1137.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology*, *5*, 169–175.

Squire, L. R., Cohen, N. J., & Nadel, L. (1984). The medial temporal region and memory consolidation: a new hypothesis. In H. Weingarter & E. Parker (Eds.), *Memory consolidation* (pp. 185–210). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*, 1380–1386.

Srinavas, K., Breedin, S. D., Coslett, H. B., & Saffran, E. M. (1997). Intact perceptual priming in a patient with damage to the anterior inferior temporal lobes. *Journal of Cognitive Neuroscience*, *9*, 490–511.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience*, *3*, 1237–1238.

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, *4*, 374–391.

Van Ormer, E. B. (1932). Retention after intervals of sleep and waking. *Archives of Psychology*, *137*, 49.

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *255*, 676–679.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*, 409–415.

APPENDIX

The model is based on binary, stochastic nodes that fire synchronously. The firing thresholds of the nodes in a module are controlled by a threshold control mechanism: inhibition in a module is diminished if there are not enough activated nodes (i.e., less than some target number k) and increased if there are too many. At each iteration, after all node activations have been updated, a learning rule is applied to all connections. The details of these mechanisms are described below. All simulations were done with the Nutshell simulator (available free of charge via http://nutshell.neuromod.org).

Activation rule

A node i has an activation a_i that can take on either of two values: 0 or 1. The probability that node i will "fire" (i.e., that its activation becomes 1) increases with its net input, as follows:

$$p_i = \frac{1}{1 + e^{\frac{net_i}{temp}}} \tag{1}$$

where net_i is the total input activation to node i, or the weighted input to node i minus inhibition:

$$\mathrm{net_i} = \sum_{j=1}^{n} w_{ij} a_j - inhibition \tag{2}$$

where w_{ij} is the connection weight from node j to node i, a_j is the activation value of node j, and n is the number of nodes in the model (if there is no connection between j and i, w_{ij} is zero by default). *Inhibition* is discussed in the next paragraph. The temperature parameter *temp* in Equation 1 controls the degree of randomness of the nodes: if *temp* is near zero, the nodes behave as simple threshold devices, if *temp* is high, the role of the net input is limited and the node takes on values 0 or 1 randomly. We used a temperature of 0.2 in all simulations.

Threshold control

The total number of activated nodes in a module (called A) is constantly monitored, and firing thresholds are adjusted to ensure that this number does not wander too far from the target number k. Each module has its own k, and inhibition control in a module is independent of that in other modules. To keep the number of activated nodes A as close as possible to the target number k, two thresholds T and τ are constantly adapted. Inhibition is the sum of a fast-changing threshold parameter T multiplied by the number of active nodes A, and a slow moving threshold τ.

$$inhibition = TA + \tau \tag{3}$$

T may reflect the excitability of the basket cells by the excitatory neurons. Slow inhibition, modelled by the threshold τ, may reflect the autonomous activity of inhibitory cells.

The control of fast inhibition, T, is straightforward: if the total activation at time t (A_t) is higher than k, T is increased (more inhibition), if A_t, is lower, it is decreased:

$$\begin{gathered} if\ A_t > (1 + crit)k \\ T = T + \Delta_t \\ if\ A_t < (1\text{-}crit)k \\ T = T - \Delta_t \end{gathered} \tag{4}$$

where *crit* is the criterion for deciding whether A_t is much larger or smaller, and Δ_t is the change made to T (*crit* = 0.20, and Δ_t = 0.01). If A is only a little bit larger or smaller than k (e.g., k < Ai < (1+crit) * k), then only one third of Δ_t was added to or subtracted from T.

One disadvantage of this method is that T may change too quickly, causing violent oscillations in activity. To prevent this, A_t is dampened by making it a moving average of the current activation and the activation of previous iterations. When $A*_t$ is the current level of activation, the value used to compute both the level of inhibition A_tT and the change in the parameter T is:

$$A_t = 0.5A_{t-1} + 0.5A*_t \tag{5}$$

This precedes calculation of the new threshold T (Eq. 4).

The slow inhibition process aims to keep the ''slow threshold'' τ equal to TA. When the equilibrium is disturbed, for example, if the activation is diminished due to a lesion, τ slowly decreases to a new equilibrium value. The speed of this change is determined by the parameter Δ_τ, which is chosen low (0.001). The expression for calculating τ_{t+1} at t+1 is

$$\tau_{t+1=(1-\Delta_\tau)\tau_t+\Delta_\tau TA} \tag{6}$$

The amount of ''fast'' inhibition is bounded by a minimum value T^{min} and a maximum value T^{max}. If $T < T^{min}$ it is set to T^{min}, and if $T > T^{max}$ it is set to T^{max}. Similarly, τ is also kept between upper and lower bounds: if $\tau < \tau^{min}$, τ is τ^{min}; if $\tau > \tau^{max}$, τ is τ^{max}. T^{min} and τ^{min} were set to 0. T^{max} *and* τ^{max} were set to such high values that they were never reached in the simulations.

Learning rule

The learning rule is a simple Hebbian rule that also allows decreases in weight. The change in weight Δw_{ij} on each time step is equal to:

$$\Delta w_{ij} = \mu^- a_i(1 - a_j), \tag{7}$$

where μ^- and μ^+ represent the learning rates. Both μ^- and μ^+ must be larger than 0. The weights w_{ij} are kept within the interval [0,1] by setting $w_{ij} = 1$ if $w_{ij} > 1$, and $w_{ij} = 0$ if $w_{il} < 0$.

Parameter settings

The values of what one could call ''equilibrium parameters'', such as the temperature, are given in the text above. The function of these parameters is to keep the number of active nodes as close as possible to the equilibrium value k, and to prevent wild swings in the number of active nodes from one iteration to the next. The remaining parameters, those that may influence results more directly than equilibrium parameters, are listed in Table 1. All parameters have the same value as was used in the simulations described in Murre and Meeter (2002).

The learning rates are given per iteration. During learning, each pattern was learned for one iteration. To limit the impact of random fluctuations in activity level on the consolidation process, consolidation is stretched out over eight iterations at a low learning rate (see Table 1). Total learning in the trace system during one consolidation period was the same as that during one acquisition period.

TABLE 1
Parameters in the connectionist simulations of the TraceLink model

Learning rate during acquisition (μ^+)	
• within trace	0.06
• within link and between the layers	0.4
Learning rate during consolidation (μ^+)	
• within trace	0.0025
• within link and between the layers	0
Total learning during one acquisition	
• within trace	0.06 (0.06 * 1 iteration)
• within link and between the layers	0.4 (0.4 * 1 iteration)
Total learning during one consolidation period	
• within trace	0.06 (0.0025 * 3 trials of 8 iterations)
• within link and between the layers	0
Unlearning rate (μ^-)	
• in all connections	75% of the learning rate
Number of nodes	
• in trace	200
• in link	42
Number of nodes active in equilibrium (*k*) ***(= number of nodes in one pattern)***	
• in trace	10 (1/20th of the nodes)
• in link	7 (1/16th of the nodes)